UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/731,929	12/10/2003	Thomas Kemp	282711US8X	6660

22850 7590 03/17/2008
OBLON, SPIVAK, MCCLELLAND MAIER & NEUSTADT, P.C.
1940 DUKE STREET
ALEXANDRIA, VA 22314

EXAMINER	
GODBOLD, DOUGLAS	

ART UNIT	PAPER NUMBER
2626	

NOTIFICATION DATE	DELIVERY MODE
03/17/2008	ELECTRONIC

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

Notice of the Office communication was sent electronically on above-indicated "Notification Date" to the following e-mail address(es):

patentdocket@oblon.com
oblonpat@oblon.com
jgardner@oblon.com

	Application No.	**Applicant(s)**
Office Action Summary	10/731,929	KEMP ET AL.
	Examiner	**Art Unit**
	DOUGLAS C. GODBOLD	2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE _3_ MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on _11 December 2007_.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) _1-9 and 12-15_ is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) _1-9, and 12-15_ is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. This Office Action is in response to communications filed December 11, 2007. Claims 1-15 are pending in the application and have been examined.

Continued Examination Under 37 CFR 1.114

2. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on December 11, 2007 has been entered.

Response to Amendment

3. The amendments filed December 11, 2007 have been considered and accepted in this office action. Claims 1, 9, and 12-14 have been amended and claim 15 has been added.

Response to Arguments

4. Applicant's arguments filed December 11, 2007 have been fully considered but they are not persuasive.

5. With regards to applicant's arguments, see pages 7-10, that Lee, Brandstein, and Gable, or any combination thereof fails to teach the new limitations "the absolute loudness being a loudness of the speech at a location of a source of the speech", the examiner respectfully disagrees. As mentioned by the applicant in the argument, most prior art systems "normalize" the input speech signal to 1, negating any factors of distance from the loudness parameters. As discussed in the last office action Brandstein discloses a relationship between the detected loudness, source distance and loudness at the source. One of ordinary skill in the art could recognize from this relationship could be used to normalize the detected loudness, instead of normalizing to 1, and that this would be merely a matter of design choice.

6. With regards to applicant's arguments, see page 10, that Lee is not combinable with Lee, the examiner respectfully disagrees. Lee was used in previous rejections to teach a method of processing speech, not to teach a phone system. Brandstein was used to teach distance location and therefore source volume location. The method of Lee could obviously be applied in a system that is not necessarily of a phone system. Therefore Lee and Brandstein are combinable arts.

Claim Rejections - 35 USC § 103

7. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

8. Claims 1-9, and 12-14 rejected under 35 U.S.C. 103(a) as being unpatentable over Lee et al. (Recognition of Negative Emotions from the Speech Signal) in view of Brandstein et al. (Microphone Array Localization Error Estimation with Application to sensor Placement).

9. Consider claim 1, Lee teaches a method for processing speech (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; page 240, column 2, lines 3-4.), comprising the steps of:

receiving a speech input of a speaker (The speech data used in the experiments was obtained from real users engaged in a spoken dialog with a machine agent over the telephone; page 241, column 1, lines 5-7.),

generating speech parameters from said speech input (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.),

determining parameters describing an absolute loudness of said speech input (The acoustic features chosen for emotion recognition comprised utterance-level statistics obtained from the pitch and energy information of the signal. These included mean median, standard deviation, maximum and minimum for energy; page 241, column 1, lines 57-61. Energy is the amplitude, and therefore the loudness of the signal.),

evaluating said speech input and/or said speech parameters using said parameters describing the absolute loudness (This paper reports on methods for

automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13.).

Lee does not specifically teach the absolute loudness being a loudness of the speech at a location of a source of the speech.

In the same field of speech processing, Brandstein suggests the absolute loudness being a loudness of the speech at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time if the invention to combine the absolute loudness as suggested by Brandstein with the speech system of Lee in order to provide a method of normalizing the loudness for emotion detection.

10. Consider claim 2, Lee teaches a method according to claim 1, wherein the step of evaluation comprises a step of emotion recognition (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; page 240, column 2, lines 3-4.).

11. Consider claim 4, Brandstein teaches a method according to claim 1 wherein a microphone array comprising a plurality of microphones (see figure 6) is used for determining said parameters describing the absolute loudness (Existing array systems have been used in a number of applications. These include teleconferencing, speech recognition, speaker identification, speech acquisition in an automobile environment, sound capture in reverberant enclosures, large room recordings, conferencing, acoustic surveillance, and hearing aid devices; page 1 lines 11-15. Obviously, the array of microphones would be used to determine the parameters including loudness needed for these applications.).

12. Consider claim 5, Brandstein teaches a method according to claim 1 wherein a location and/or distance of the speaker is determined (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5.).

13. Consider claim 6, Brandstein teaches a method according to claim 1 wherein the absolute loudness is determined using algorithms for auditory and/or binaural processing (Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array.).

14. Consider claim 7, Brandstein teaches a method according to claim 5, wherein said absolute loudness is computed by normalizing a measured loudness, or energy by said distance (Page 21 provides a relationship of a source amplitude as a function of distance and angle form the source. This relationship could obviously be used to normalize an amplitude value to estimate the amplitude at the source.)

15. Consider claim 8, Brandstein teaches a method according to claim 5, wherein said distance is determined using the time delay of the speech input between said plurality of microphones (Sections 2 and 3 discuss using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-10.)

16. Consider claim 9, Lee teaches a speech processing system (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; page 240, column 2, lines 3-4.), configured to:

receive a speech input of a speaker (The speech data used in the experiments was obtained from real users engaged in a spoken dialog with a machine agent over the telephone; page 241, column 1, lines 5-7.),

generate speech parameters from said speech input (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.),

determine parameters describing an absolute loudness of said speech input (The acoustic features chosen for emotion recognition comprised utterance-level statistics

obtained from the pitch and energy information of the signal. These included mean median, standard deviation, maximum and minimum for energy; page 241, column 1, lines 57-61. Energy is based the amplitude, and therefore the loudness of the signal.),

evaluate said speech input and/or said speech parameters using said parameters describing the absolute loudness (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13.).

Lee does not specifically teach the absolute loudness being a loudness of the speech at a location of a source of the speech.

In the same field of speech processing, Brandstein suggests the absolute loudness being a loudness of the speech at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time if the invention to combine the absolute loudness as suggested by Brandstein with the speech system of Lee in order to provide a method of normalizing the loudness for emotion detection.

17. Consider claim 12, Lee teaches a computer readable medium encoded with a computer program configure to cause a processor based device to execute the method of: (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; page 240, column 2, lines 3-4. a computer readable medium is inherent as this is computer based.):

receiving a speech input of a speaker (The speech data used in the experiments was obtained from real users engaged in a spoken dialog with a machine agent over the telephone; page 241, column 1, lines 5-7.),

generating speech parameters from said speech input (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.),

determining parameters describing an absolute loudness of said speech input (The acoustic features chosen for emotion recognition comprised utterance-level statistics obtained from the pitch and energy information of the signal. These included mean median, standard deviation, maximum and minimum for energy; page 241, column 1, lines 57-61. Energy is the amplitude, and therefore the loudness of the signal.),

evaluating said speech input and/or said speech parameters using said parameters describing the absolute loudness (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13.).

Lee does not specifically teach the absolute loudness being a loudness of the speech at a location of a source of the speech.

In the same field of speech processing, Brandstein suggests the absolute loudness being a loudness of the speech at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time if the invention to combine the absolute loudness as suggested by Brandstein with the speech system of Lee in order to provide a method of normalizing the loudness for emotion detection.

18. Consider claim 13, Lee teaches a method for processing speech, comprising:

receiving a speech signal of a speaker (The speech data used in the experiments was obtained from real users engaged in a spoken dialog with a machine agent over the telephone; page 241, column 1, lines 5-7.);

generating speech parameters from said speech signal (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.); and

evaluating at least one of said speech signal and said speech parameters using the normalized loudness or energy (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13.).

However Lee does not specifically teach:

determining a distance of the speaker based on a time delay of a respective arrival of said speech signal at two or more microphones; and

normalizing a measured loudness or energy by said distance, and

calculating an absolute loudness being a loudness of a speech that generated the speech signal at a location of a source of the speech.

In the same field of speech processing, Brandstein teaches determining a distance of the speaker based on a time delay of a respective arrival of said speech signal at two or more microphones (Sections 2 and 3 discuss using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-10.); and

normalizing a measured loudness or energy by said distance (Page 21 provides a relationship of a source amplitude as a function of distance and angle form the source. Although this relationship was given to model the source, one of ordinary skill in the art at the time of the invention would have thought, given the location of the source (as determined in the localization method discussed throughout Brandstein) and the detected amplitude at the microphone array, to use the relationship to determine the source amplitude) and

calculating an absolute loudness being a loudness of a speech that generated the speech signal at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a microphone array for source location and absolute volume as suggested by Brandstein with the speech processing system of Lee in order to provide a means for provide a high quality signal of the desired speaker that is not adversely effected by the distance from a speaker to the microphone array. (Introduction, Brandstein.).

19. Consider claim 14, Lee teaches a system for emotion recognition and/or speaker identification, comprising:

a data processor configured to generate speech parameters from said speech signal (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.), and

further configured to evaluate at least one of said speech signal and said speech parameters using the normalized loudness or energy (This paper reports on methods

for automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13.).

However Lee does not specifically teach:

at least two microphones configured to receive a speech signal; and

a processor configured to determine a distance of the speaker based on a time delay of a respective arrival of said speech signal at said microphone, to normalize a measured loudness or energy by said distance and calculating an absolute loudness being a loudness of a speech that generated the speech signal at a location of a source of the speech

In the same field of speech processing Brandstein teaches:

at least two microphones configured to receive a speech signal (see microphone array in figure 6); and

a processor configured to determine a distance of the speaker based on a time delay of a respective arrival of said speech signal at said microphone (Sections 2 and 3 discuss using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-10.), to normalize a measured loudness or energy by said distance (Page 21 provides a relationship of a source amplitude as a function of distance and angle form the source. Although this relationship was given to model the source, one of ordinary skill in the art at the time of the invention would have thought, given the location of the source (as determined in the localization method discussed throughout Brandstein) and the detected amplitude at the microphone array, to use the relationship to determine the source amplitude) and calculating an absolute loudness

being a loudness of a speech that generated the speech signal at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to use a microphone array for source location and absolute volume as suggested by Brandstein with the speech processing system of Lee in order to provide a means for provide a high quality signal of the desired speaker that is not adversely effected by the distance from a speaker to the microphone array. (Introduction, Brandstein.).

20. Consider claim 15, Lee teaches a method for processing speech (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; page 240, column 2, lines 3-4.) comprising the steps of:

receiving a speech signal of a speaker (The speech data used in the experiments was obtained from real users engaged in a spoken dialog with a machine agent over the telephone; page 241, column 1, lines 5-7.);

calculating an absolute loudness (The acoustic features chosen for emotion recognition comprised utterance-level statistics obtained from the pitch and energy

information of the signal. These included mean median, standard deviation, maximum and minimum for energy; page 241, column 1, lines 57-61. Energy is the amplitude, and therefore the loudness of the signal.);

determining features from the speech signal, wherein the features are at least partly based on the absolute loudness (In our experiments, we computed only acoustic features such as pitch and energy related features from the speech signal; page 241, column 2, lines 46-47.); and

determining an emotion and/or an identity of the speaker based on the features (This paper reports on methods for automatic classification of spoken utterances based on the emotional state of the speaker; using utterance level features; page 240, column 2, lines 3-13).

Lee does not specifically teach the absolute loudness being a loudness of the speech at a location of a source of the speech.

In the same field of speech processing, Brandstein suggests the absolute loudness being a loudness of the speech at a location of a source of the speech (Section 2 discusses using a microphone array with a time difference of arrival algorithm to determine a location of a speaker; pages 3-5. Page 21 teaches modeling a source as a cardioid radiator, wherein the source amplitude is a function of distance from the source. When this information is combined with the source locating algorithms of section 2, one can obviously estimate the amplitude at the source itself given the amplitude at the microphone array).

Therefore it would have been obvious to one of ordinary skill in the art at the time if the invention to combine the absolute loudness as suggested by Brandstein with the speech system of Lee in order to provide a method of normalizing the loudness for emotion detection.

21. Claim 3 is rejected under 35 U.S.C. 103(a) as being unpatentable over Lee in view of Brandstein as applied to claim 1 above and further in view of Gable et al. (US PAP 2005/0060153).

22. Consider claim 3, Lee and Brandstein teaches the method according to claim 1 but does not specifically teach wherein the step of evaluation comprises a step of speaker identification.

In the same field of speech processing, Gable teaches a step of speaker identification using similar acoustic features as described by Lee (Verification parameters represent the individuality of the speaker, containing information about the timing, pitch, amplitude or spectral content of the speech; paragraph 0027. Abstract discusses using these features for speaker verification.).

Therefore it would have been obvious to one of ordinary skill in the art at the time of the invention to provide speaker identification as taught by Gable, with the speech processing of Lee in order to provide a method of further classifying a speech signal beyond emotional classification.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to DOUGLAS C. GODBOLD whose telephone number is (571)270-1451. The examiner can normally be reached on Monday-Thursday 7:00am-4:30pm Friday 7:00am-3:30pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Patrick Edouard can be reached on (571) 272-7603. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

DCG

/Talivaldis Ivars Smits/
Primary Examiner, Art Unit 2626